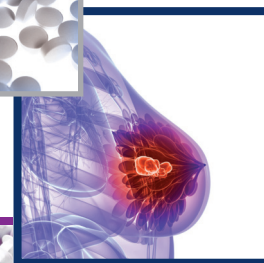


Real and artificial controversies in breast cancer screening

Stephen W Duffy*¹, Tony Hsiu-Hsi Chen², Robert A Smith³, Amy Ming-Fang Yen⁴ & Laszlo Tabar⁵



Practice Points

- Mammographic screening prevents deaths from breast cancer and can be recommended.
- When expressed in terms of breast cancer mortality in women of average UK risk, screened regularly from the age of 50–69 years, the major reviews indicate a reduction in breast cancer mortality in the range of one life saved per 64–257 women screened.
- Offering screening on the basis of age (starting at either 40 or 50 years) is an effective strategy.
- There is scope for varying the surveillance regimen (including the imaging technology) based on family history or breast density.

SUMMARY We review the apparent disparities between different reviews of the effects of mammographic screening on mortality from breast cancer and overdiagnosis. When results of each review are expressed with respect to a common population and a common baseline, all find a substantial mortality benefit and variation among estimates is minor. There are genuine disagreements about overdiagnosis, but methods that take account of lead time and underlying incidence trends yield estimates of overdiagnosis that are modest and are outweighed by the mortality benefit. There is potential for individualized screening regimens, particularly with respect to breast density.

In this paper, we give our perspective on recent debates about breast cancer screening, based on our own experience and knowledge of the research literature. It is generally accepted that screening with mammography prevents deaths from breast cancer, although debate continues about the absolute size of the mortality benefit conferred and the concomitant risks associated with screening [1–3]. Among these risks, the most highly publicized in recent years is overdiagnosis, defined as the diagnosis by screening

of cancer that would not have been diagnosed in the patient's lifetime if screening had not taken place [4]. A number of recent observational studies have claimed to find low rates of benefit in terms of reducing mortality rates or late stage disease, and high rates of overdiagnosis [5,6]. These have achieved a high profile in the mass media and stimulated further debate [7,8]. Other areas of debate include ages at which to begin and end screening, the interaction of screening and contemporaneous changes to therapy, and

¹Centre for Cancer Prevention, Wolfson Institute of Preventive Medicine, Barts & The London School of Medicine & Dentistry, Queen Mary University of London, Charterhouse Square, London, EC1M 6BQ, UK

²Graduate Institute of Epidemiology & Preventive Medicine, National Taiwan University, Taipei, Taiwan

³Cancer Control Science Department, American Cancer Society, Atlanta, GA, USA

⁴School of Oral Hygiene, Taipei Medical University, Taiwan

⁵Department of Mammography, Falun Central Hospital, Sweden

*Author for correspondence: Tel.: +44 20 7882 3535; Fax: +44 20 7882 3890; s.w.duffy@qmul.ac.uk

the potential for replacing general screening recommendations based on age with individually tailored risk-based screening [9–12].

While it may be worth revisiting the substantive arguments on these issues, it would seem more profitable first to establish whether the conflicting claims for reductions in mortality and rates of overdiagnosis represent genuine quantitative and qualitative disagreements or simply stem from different modes of expression of the same results. The recent UK review of mammographic screening correctly reiterated the key methodological point that the estimated absolute numbers of deaths prevented will depend on the period of observation and the population to which screening is applied, and that estimates of overdiagnosis will differ according to the denominator relative to which overdiagnosis is expressed [2]. Therefore, very different numerical estimates are derived depending on whether overdiagnosis is expressed as a population rate, as a proportion of lifetime risk of breast cancer or as a proportion of screen-detected cancers. Thus, before identifying which estimates are most likely to be correct, it would be valuable to assess the extent to which they are different ways of expressing the same results.

In this paper, we aim to examine: the magnitude and source of genuine disagreement over the absolute mortality benefit of breast cancer screening, including the issue of the role for early detection in the epoch of adjuvant systemic therapies; the magnitude and source of genuine disagreement over the rate of overdiagnosis conferred by screening; and the appropriate populations to be offered screening.

The reduction in breast cancer mortality

The randomized trials of mammographic screening have been reviewed many times; indeed there have been many more reviews of the trials than trials themselves, and considerable debate on the relative merits of the various reviews. When synthesizing the results of the trials in terms of relative reduction in the risk of dying from breast cancer associated with an invitation to screening, reviewers invariably arrive at an estimated breast cancer mortality reduction in the region of 20% [2,3,9,13]. Some find a slightly lower benefit due to use of particular follow-up periods or exclusion of certain studies [9], others a slightly higher benefit from using only the population-based trials [13], but the relative benefit remains close to 20%. One exception is the Nordic Cochrane

review, which discards the empirical results in favor of a conjectured mortality reduction of 15% on the basis of the reviewers' judgment of the quality of the trials reviewed [3].

Despite the similarity in relative benefit estimates, the absolute reductions in mortality differ widely among reviewers [2,3,9,13,14]. The absolute benefit is often expressed as number needed to screen to prevent one breast cancer death, by analogy with the number needed to treat often reported in the results of therapeutic trials. It has been argued that the measure is not appropriate to primary or secondary prevention [15], but it is in widespread use. In the case of mammography screening, it is complicated by the fact that some reviewers report number needed to invite rather than number needed to screen [3,9]. This is not a useful measure, since the number needed to invite will vary depending on the proportion of invitees who take up the offer of screening, which can be highly variable, and since invitation of itself does not confer any protection [16]. The number needed to invite will be biased by deaths in nonattenders, to an extent proportional to the nonattendance rate in the specific invited population.

Table 1 shows the absolute benefit figures quoted by the four highest profile recent reviews, which represent the range of estimates of the breast cancer mortality reduction. These are the UK Independent Review, the Nordic Cochrane review, the US Preventive Services Task Force (USPSTF) review and the EUROSCREEN review of mammography service screening in Europe [2,3,9,14]. The first three are reviews of the randomized trials of screening, the fourth, EUROSCREEN, a review of observational estimates pertaining to service screening in Europe. The estimated number needed to screen/invite to prevent one death from breast cancer ranges from 111 to 2000, almost a 20-fold range. The question arises: why is this range so great when the relative benefit estimates only vary by a factor of two to three? The estimated number required varies by age group, whether the intervention described is actual screening or invitation to screening, follow-up time and other factors.

To assess whether these represent genuine disagreements or whether the differences are mainly due to such factors as follow-up time and target population, we converted all four to pertain to the same scenario as used in the UK Independent Review, that is to the effect of screening for 20 years from the age of 50–69 years on breast

cancer mortality at age 55–79 years, in a UK population. In this, no recalculation is required for the UK Independent Review.

The Nordic Cochrane review claims a 15% reduction in breast cancer mortality, pertaining to invitation to screening. In the UK, cumulative mortality from breast cancer over ages 55–79 years with the current screening program is 17 per 1000 [2]. From this, the mortality in the absence of screening would be expected to be $17/0.85 = 20$ per 1000. The effect of invitation would therefore be three breast cancer deaths prevented per 1000 invited. With 77% attendance for screening, the corresponding estimated effect of being screened is $3/0.77 = 3.89$ deaths per 1000 invited, and $1000/3.89 = 257$ needed to screen to prevent one breast cancer death. The major reasons for the Cochrane authors' very low estimate of absolute benefit are the restriction of benefit to a ten-year period and the estimation of absolute mortality rates from a selected minority of trials dominated by the age group 40–49 years, which has considerably lower absolute mortality from breast cancer than the typical age range offered screening.

The USPSTF relative risk estimates associated with invitation to screening were 0.86 (95% CI: 0.75–0.99) and 0.68 (95% CI: 0.54–0.87) for ages 50–59 and 60–69 years, respectively. Taking an inverse-variance weighted average of these in the logarithmic scale gives an overall estimate of 0.81 (95% CI: 0.72–0.92) for ages 50–69 years, similar to the previously described meta-analysis results. As noted above, in the UK, cumulative mortality from breast cancer over ages 55–79 years with the current screening program is 17 per 1000 [2]. From the USPSTF estimate, we would expect the mortality in the absence of screening to be $17/0.81 = 21$ per 1000. Thus, the estimated effect of the policy of offering screening is to prevent four deaths per 1000 women invited. With 77% attendance for screening, this means that $4/0.77 = 5.19$ deaths per 1000 screened are estimated to be prevented over the 25-year period on the basis of the USPSTF results. This translates to 193 needed to screen to prevent one breast cancer death.

The EUROSCREEN estimated reductions in breast cancer mortality associated with screening were 38% for the incidence-based mortality studies and 48% for the case–control studies. For the former figure, applying it to the UK rate of 17 per 1000 would give 27.4 deaths per 1000 expected in the absence of screening, such as

Table 1. Estimates of mortality reductions associated with mammographic screening or invitation to screening from four major reviews.

Review	Breast cancer mortality reduction (%)	Intervention	Source of absolute mortality rate	Screening period, ages (years)	Follow-up period, ages (years)	NNS	Ref.
UK independent review	20	Screening	National rates, UK	20, ages 50–69	25, ages 55–79	180	[2]
Nordic Cochrane Review	15	Invitation	3 trials dominated by ages 40–49 years	10 years, during ages 40–74	10, during ages 40–74	2000	[3]
USPSTF, ages 39–49 years	15	Invitation	Trials in review	Average 7, beginning at age 39–49	Average 14, beginning at age 39–49	1904	[9]
USPSTF, ages 50–59 years	14	Invitation	Trials in review	Average 7, beginning at age 50–59	Average 14, beginning at age 50–59	1339	[9]
USPSTF, ages 60–69 years	32	Invitation	Trials in review	Average 7, beginning at age 60–69	Average 14, beginning at age 60–69	337	[9]
EUROSCREEN [†]	38–48	Screening	National rates, UK, Nordic countries and Italy, ages 50–79 years	20 years, ages 50–69	30 years, ages 50–79	111–143	[14]

[†]Percent mortality reductions are based on invitation to screening, not actually attending screening.

[‡]Prospective incidence-based mortality studies found the lower reduction, case–control studies the higher.

NNS: Number needed to screen or invite to prevent one breast cancer death; USPSTF: US Preventive Services Task Force.

5.6 deaths prevented per 1000 and 96 needed to screen to prevent one breast cancer death. The same calculations for the case-control estimate would give 64 needed to screen.

Table 2 shows the results for all four reviews converted to the UK review scenario. As can be seen, the range of absolute benefits is now only four- rather than 20-fold. The differences between the estimates from reviews of the randomized trials is less than 1.5-fold. Even the Nordic Cochrane Review estimate is in the same region as the other estimates. Thus, the differences between the reviews with respect to the absolute breast cancer mortality reduction are almost entirely due to expressing the same basic result relative to different denominators, choice of population mortality rates and so on. Thus, the so-called controversy over the benefit of mammography screening as estimated from the trials is largely artificial. When expressed relative to the same denominator, with the same screening and follow-up periods, and using the same absolute mortality rates, absolute benefit estimates are all of the same order of magnitude (and all indicate a substantial reduction in breast cancer mortality with screening).

Some have raised the question of whether the benefits of screening observed in the randomized trials, which took place largely before 1990, will still be observed in the epoch of systemic adjuvant therapies and improved standardization of management of breast cancer generally. That is, do new treatments render early detection unnecessary? In one sense, this might be considered the wrong issue, in that one might argue that since diagnosis precedes treatment, it makes more sense to ask if screening renders some of the therapeutic developments redundant. However, whether or not advances in therapy have diminished the importance of screening is a fair question, so next we will consider prognosis of breast cancers diagnosed in the modern therapeutic epoch.

Figure 1 shows survival by node status in 9040 patients with breast cancers diagnosed in the East of England in 1998–2003. Clearly, even in the era of standardized treatment and adjuvant systemic therapy, node-negative cancers still have considerably better survival than node positive. To avoid issues of lead time bias, Figure 2 shows the corresponding survival for the subset of 6875 symptomatic tumors. Again, there is a clear survival advantage of node-negative disease. A similar relative benefit (albeit with better absolute survival) can be seen for the 2165 screen-detected tumors only (Figure 3). Thus, it is clear that improvements in therapy have not rendered early detection redundant. While survival of women with node-positive tumors has improved in the era of adjuvant systemic therapy compared with earlier times, it clearly is still preferable and therefore a worthy goal to have the cancer diagnosed before it spreads to the lymph nodes. These survival figures themselves do not demonstrate the benefit of screening. That is the role of the randomized trials. Nor do they deny improvements in treatment. They do show, however, that even in the current therapeutic environment there is still a substantial benefit of earlier diagnosis, while the tumor is node negative.

To conclude, the disagreement over the absolute mortality benefit of screening as estimated from the randomized controlled trials is largely artificial, due mainly to different modes of expression of the same essential results. The question of whether early detection is still relevant in the current therapeutic environment is a reasonable one, but the answer is clearly yes.

As noted above, in recent years, there have been a number of high-profile publications reporting analyses of published breast cancer mortality and incidence rates, claiming that little or no benefit has accrued from screening [5,6,17]. Partly due to the prestige of the journals publishing these reports, these publications have received

Table 2. Estimates of the number needed to screen (not invite) between ages 50 and 69 years, to prevent one breast cancer death at the age of 55–79 years, using the estimates from four major reviews of mammographic screening and the breast cancer mortality rates in the UK population at the age of 55–79 years.

Review	NNS	Ref.
UK Independent review	180	[2]
Nordic Cochrane review	257	[3]
USPSTF	193	[9]
EUROSCREEN [†]	64–96	[14]

[†]Prospective incidence-based mortality studies found the higher number, case-control studies the lower. NNS: Number needed to screen or invite to prevent one breast cancer death; USPSTF: US Preventive Services Task Force.

considerable attention from the mass media and from policy makers. However, reviews show that such negative findings constitute a small minority of the totality of evidence [18,19]. For purposes of screening policy and public information, it is this totality of evidence that is important.

With respect to the observational studies, they are not all of equal quality, but there is an additional issue worth highlighting. Observational studies can be useful for determining how well screening is achieving its aims now, and they may complement the randomized trial results or qualify our interpretation of these. They have been useful in addressing tactical issues in screening, such as the age ranges and screening intervals in cervical screening [20]. What they cannot generally do is outweigh the experimental evidence. Consider an example in therapeutic oncology: randomized trials show that treatment with rituximab is effective in improving survival in follicular lymphoma [21]. In the UK, male age-standardized mortality rates from lymphoma have increased from approximately four per 100,000 in the 1970s to more than six per 100,000 since 2000 [101]. An assertion that can be concluded from this is that the rituximab treatment is killing patients, and that the randomized evidence must be wrong, would not be taken seriously, and less superficial analysis of the statistics would conclude that the increased mortality is largely driven by increased incidence. The logic that the results of randomized trials take precedence over ecological findings prevails in the area of therapy. It should also do so in the field of screening and prevention.

Overdiagnosis

In the last decade, some authors have inferred very high rates of overdiagnosis from increasing incidence rates of breast cancer [5,22,23]. These estimates have for the most part been derived with no information on which individuals were screened or which cancers were screen detected, which logically is a necessary underpinning of the estimate if overdiagnosis occurs as a result of exposure to screening. To the extent that some overdiagnosis of breast cancer may occur, to be reliable, estimates of overdiagnosis derived from breast cancer incidence rates over time should take account of any long-term underlying trends in incidence taking place independent of screening, and of the phenomenon of lead time – that is, the estimate of overdiagnosis should not include cancers detected a few years earlier than they

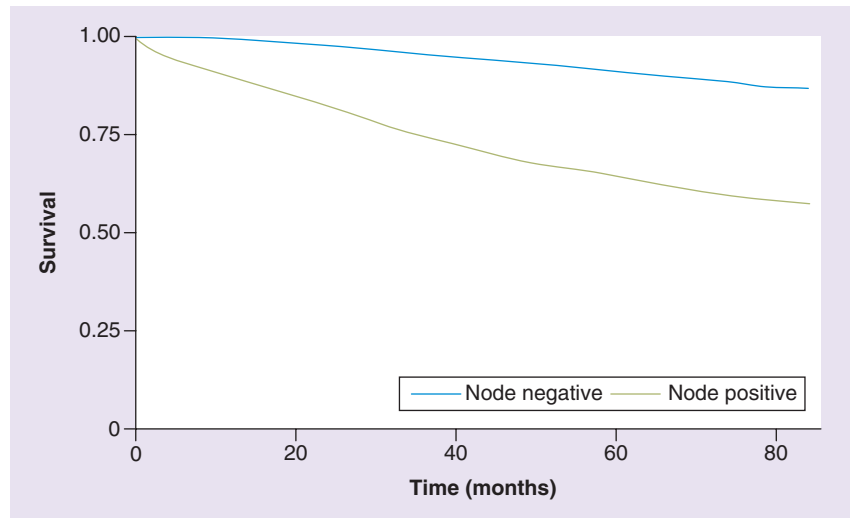


Figure 1. Survival by node status in 9040 breast cancers diagnosed between 1998 and 2003.

would have been in the absence of screening. It has been pointed out that the most recent and arguably the highest profile publication, that of Bleyer and Welch [5], fails to take full account of underlying incidence trends [8], that the publication by Kalager *et al.* [23] fails to fully adjust for lead time [24] and the study of Jørgensen *et al.* [22] has been criticized on both counts [25].

Puliti *et al.* have shown that studies that fail to take full account of lead time and of underlying incidence trends tend to estimate high rates of overdiagnosis of the order of 30–50% [4]. Those that do take account of these phenomena obtain much lower rates, of the order of 10% or less. Taking account of lead time requires either very long follow-up or use of external estimates of the preclinical detectable period. Correction for underlying incidence trends may involve extrapolation of prescreening trends. Thus, correction can never be perfect. However, ignoring these phenomena results in a higher cost to validity than estimating them. Duffy and Parmar [24] show that lead time alone can generate artificial increases in incidence of the order of 30–50%, and that failure to include data up to 10 years after the age that screening ceases will cause overdiagnosis to be overestimated. This suggests that almost all published estimates of overdiagnosis, including our own, exaggerate the problem of excess incidence attributable to screening [2–5,22,23]. Thus, while it appears that there is genuine disagreement over the extent of overdiagnosis in breast cancer screening the case for high rates of overdiagnosis rests on analyses

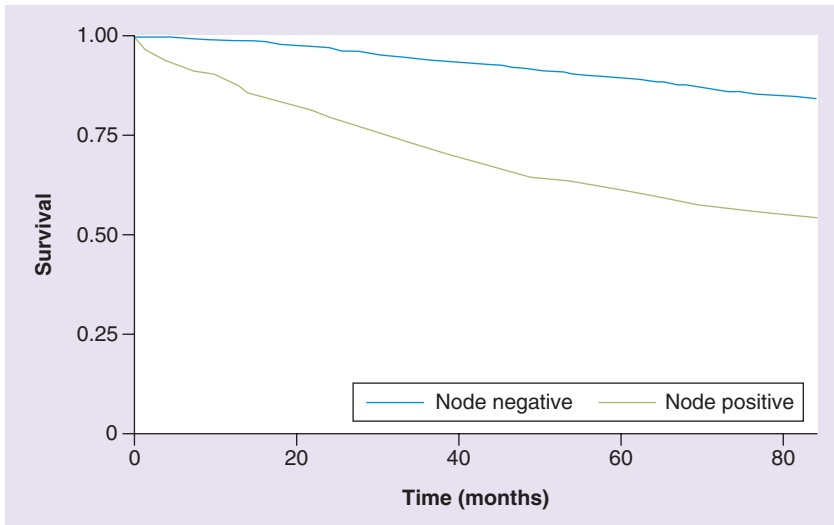


Figure 2. Survival by node status in 6875 symptomatic breast cancers diagnosed between 1998 and 2003.

that are biased by lead time, and incidence trends occurring independently of screening. It may be possible to transform reported estimates of overdiagnosis to a standard absolute estimate as we have done for the mortality reductions in Table 2, but this would be a major methodological exercise, and beyond the scope of this paper.

The regular publication of overdiagnosis estimates that do not fully adjust for such well-known complicating factors such as underlying trends in incidence and lead time is perplexing. It may be that there is reluctance on the part of peer-reviewers and editors to disagree with what might be considered healthy skepticism. This may be seen in the publication of Jørgensen and colleagues' conclusion from published incidence figures that the UK screening program has caused overdiagnosis of the order of 30–40% [22]. This was published despite a number of questionable methodological approaches, including:

- Estimation of incidence by eye from a published graph;
- Use of linear regression rather than Poisson regression to analyze rate data;
- Exclusion of the 3 years of highest incidence in the prescreening period, and inclusion of only the year of highest incidence in their post-screening epoch;
- Inadequate adjustment for prescreening trends, as evidenced by the 7% excess incidence observed in the screening epoch in the

under-50 years age group, which was not exposed to screening;

- Lack of correction for lead time;
- Use of assumptions in the absence of data on incidence and overdiagnosis of ductal carcinoma *in situ*.

The point of this is to question why these methodological shortcomings were not dealt with at an earlier stage. The conclusions from the discussion above should be twofold. Firstly, that overdiagnosis is quantitatively minor compared to the breast cancer mortality reduction, as judged by the evidence that takes into account the effects of lead time and underlying incidence trends. Secondly, there is a need for editors and peer-reviewers to subject skeptical views on mammography to the same critical scrutiny as favorable findings.

Who should be screened & with what regimen?

In the past, the major point of debate over mammography screening was whether or not to offer the examination to women aged 40–49 years [26]. Arguments against offering screening to this age group included the lesser relative reduction in breast cancer mortality observed in the trials in this age group, the lower incidence at ages 40–49 years compared with women aged 50 years and over, and the comparatively lower efficiency of screening women in this age group due to higher levels of breast density at younger ages [27]. However, mature follow-up of the randomized trials indicates an unequivocal breast cancer mortality reduction with the offer of screening in this age group [28]. Furthermore, in Sweden, at the time of introduction of nationwide mammography screening, the policy makers chose age 40 years as the lower age limit in approximately half of the counties in the country, and 50 in the remaining half. At 16-year observation, mortality from breast cancers diagnosed at ages 40–49 years was significantly lower in those counties that offered screening starting from the age of 40 years [10].

With regards to the lower incidence of breast cancer in women aged 40–49 years, it is worth noting that ovarian cancer screening is under trial in women aged 50–74 years [29], and incidence of ovarian cancer in this group is considerably lower than that of breast cancer at ages 40–49 years [30]. Indeed at considerably more

than 100 per 100,000 per year, breast cancer incidence is relatively high in this age group. Furthermore, other measures of disease burden beyond incidence are relevant in the decision to invite an age group to screening. Women who were diagnosed with breast cancer in their 40s account for a significant fraction of premature mortality attributable to death from breast cancer, and a similar fraction of the deaths that occur each year when compared with women aged 50–59 and 60–69 years [31].

It is true that breast density is higher in younger women [32], but the mortality reductions already noted [10,28] indicate that mammography achieves early detection and prevention of breast cancer deaths despite this, and options for supplemental imaging in the presence of significant breast density are increasing [33,34]. Thus, there is no particular reason to exclude this age group from a screening program.

The issue of density raises a highly topical issue, that of risk-based or otherwise individually tailored screening. In one sense, risk-based screening for breast cancer is already practiced, in that target populations are selected on the basis of two of the strongest risk factors for the disease – age and sex. In addition, many healthcare providers offer more intensive surveillance to women with a significant family history of breast cancer [35]. It has been suggested that basing eligibility on a combination of polygenic risk and age, and other breast cancer risk factors would improve the cost–effectiveness of screening [12,36]. This would not necessarily be an unqualified benefit. The improved cost–effectiveness would be modest and the increased complexity of the eligibility criteria might detract from the performance of the program as a whole. A public health intervention is more likely to be successful if it has a simple and transparent system of eligibility and delivery. In addition, equity issues arise, as the population excluded from screening would not be at negligible risk. While risk-based screening of the kind proposed [12,36] remains an interesting research issue, it is unlikely to find its way into practice until the risk stratification can identify a population with extremely low risk of breast cancer.

There remains the possibility of individualized screening regimens. Already, some providers offer different screening frequencies depending on age [10]. The Swedish National Board of Health and Welfare recommends mammography screening with a 12–18-month

interval for women aged 40–54 years and with an 18–24-month interval for women aged 55–74 years. There is evidence that mammographic density, in addition to being positively associated with risk and negatively associated with screening sensitivity, is also inversely related to the potential lead time [37]. The possibility of determining screening frequency on the basis of density has been suggested [38]. However, caution has been urged about such a strategy in view of our limited understanding of the biological mechanisms of the density/risk association [39]. A more direct approach to the issue of density would be to vary the screening technology to obtain sensitivity comparable to that obtained by mammography in nondense breasts, for example offering combined mammography and tomosonography to women with very dense breast tissue [34]. As digital mammography replaces film, and as our understanding of the role of breast density improves, such options will become more attractive.

The discussion above does not include at what age to stop screening. Healthcare providers have difficult decisions to make in this area. It may be that the most reasonable strategy is to stop inviting women at a certain age, but permit self-referral for screening thereafter, as in the UK program, where women aged over 70 years will be screened if they request it. Also, it should be noted that age-specific and risk factor-specific estimates of overdiagnosis are not available at the moment. This should be a target for future research.

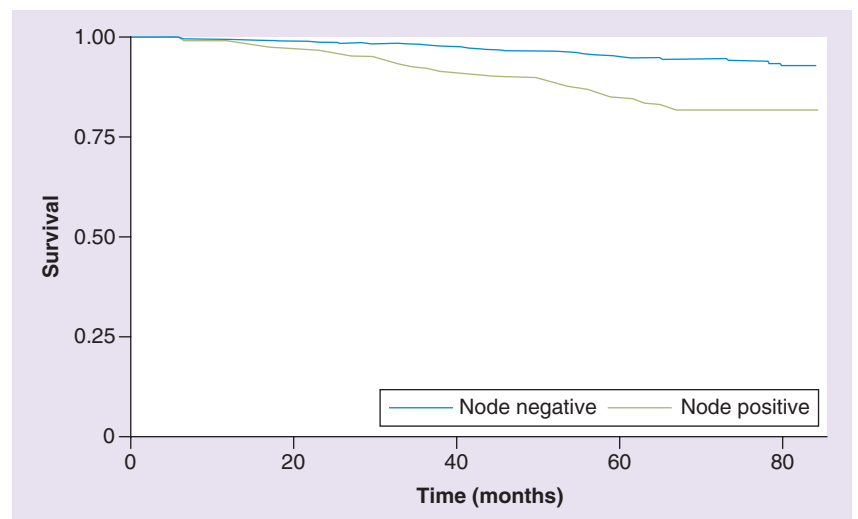


Figure 3. Survival by node status in 2165 screen-detected breast cancers diagnosed between 1998 and 2003.

Conclusion

The controversy over the effect of mammographic screening on breast cancer mortality is largely artificial. When like is compared with like, the estimates from all major reviews of the subject point to a substantial and significant reduction in breast cancer deaths with the offer of screening. There are genuine disagreements about overdiagnosis, but estimation methods that properly take into account the complicating factors of lead time and underlying incidence trends yield modest estimates of overdiagnosis. There are also varying views on the possibilities of individualized screening.

Future perspective

Physicians need to understand and be able to explain: how and why screening has successfully reduced the rate of advanced cancers and breast cancer mortality through early detection; the extent to which this can be accomplished in a modern screening program; and the impact of attending screening regularly compared with not attending. A good grasp of this evidence is the foundation for the physicians' key role in informing women about what to expect from screening, and the benefits of attending screening compared with not attending screening.

The population-based randomized controlled trials carried out in the 1970s–1990s tested and confirmed the beneficial effect of early detection and treatment of breast cancer in its early stages on mortality from the disease. Having demonstrated the efficacy of intervention early in the natural history of the disease, it is acceptable to infer that any other imaging modality in the screening setting that succeeds in significantly improving the prognostic attributes of breast cancer (tumor size, axillary node status and histologic malignancy grade) will also result in a significant decrease in the advanced breast cancer rate with a subsequent significant decrease in mortality. It is important to acknowledge that not all women will benefit equally from mammography screening, and here we particularly are referring to women with significant breast density. MRI of the breasts is already being used in addition to mammography in young women at known or suspected inherited susceptibility to breast cancer due to carrying a mutation on a breast cancer susceptibility gene [40]. MRI, the recently developed 3D automated ultrasound for women with dense breasts, and other evolving technologies could be used in the future as either an additional

screening method following mammography in women with dense breasts or as independent tools for regular surveillance of high-risk populations. Using improved cancer detection methodology to modify future screening policies will further improve the efficacy of screening.

It is a common misunderstanding that the randomized screening trials tested the impact of screening on mortality from breast cancer. In reality, it was the 'invitation to screening' that was tested in these scientific trials. Distinguishing the impact of regular attendance at screening, as opposed to being invited, will provide objective and evidence-based information to women about the impact of 'attending screening regularly'.

It is unfortunate that antiscreening reports that receive media attention have exaggerated the magnitude of the potential harms of mammography screening while simultaneously seriously underestimating the proven benefits of screening. While it is important to discuss all aspects of screening asymptomatic women, potential harms included, the harms of not attending screening are not commonly discussed. According to a recent study, women not attending screening had significantly larger tumors, worse stage at diagnosis, and worse overall and disease-specific survival [41]. This itself does not prove that screening reduces mortality from breast cancer: that was done by the randomized trials. It does, however, illustrate the consequences of being screened and of not being screened in the minority who develop breast cancer.

In conclusion, the oft-repeated statement 'mammography screening is controversial' is misleading, since it fails to distinguish between genuine areas of uncertainty that require further data, evaluation and thoughtful discussion versus those enduring issues that can fairly be described as ideological and characterized by entrenched positions. For the casual observer, it is difficult if not impossible to assess which among the polar positions is most correct. But while some controversies are genuine, others are not. As noted above, when the various review estimates of absolute benefit are expressed as number needed to screen (not invite) and applied to the same population, the disagreement among estimates is minor. One might describe overdiagnosis as controversial, but it should be borne in mind that the excessively high estimates of overdiagnosis all come from studies that typically do not have individual data on screening exposure, and which do not adequately account for trends in disease

incidence and lead time of screening. Among studies that do have individual data and/or take into account the complexities of underlying incidence trends and lead time, there is a consensus that overdiagnosis lies in the range of 0–10%, and that the benefits of screening in terms of mortality outweigh this side effect [14,42].

Mammography screening provides women with an opportunity to reduce their risk of being diagnosed with an advanced breast cancer and of dying from breast cancer [1,2,26,28,43,44]. The randomized controlled trials have demonstrated the efficacy of offering breast cancer screening, and the majority of observational studies of modern service screening have demonstrated population-based results as good as or better than those demonstrated by the randomized trials. The evidence clearly shows that benefits outweigh harms, and taken together, all but one systematic review over three decades has reached this same conclusion. It is important that healthcare professionals appreciate that the overwhelming weight of the historical and modern evidence supports the importance of routine breast cancer screening

so that they are in a position to provide the best advice to their patients.

Acknowledgements

The authors wish to thank D Greenberg of the Eastern Cancer Registration and Information Centre for the survival information in Figures 1–3.

Financial & competing interests disclosure

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

No writing assistance was utilized in the production of this manuscript.

Open access

This work is licensed under the Creative Commons Attribution-NonCommercial 3.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/3.0/>

References

Papers of special note have been highlighted as:

- of interest
 - of considerable interest
- 1 Tabar L, Vitak B, Chen TH *et al.* Swedish two-county trial: impact of mammographic screening on breast cancer mortality during 3 decades. *Radiology* 260, 658–663 (2012).
 - **Longer follow-up than any other breast screening trial.**
 - 2 Independent UK Panel on Breast Cancer Screening. The benefits and harms of breast cancer screening: an independent review. *Lancet* 380, 1778–1786 (2012).
 - **Highly influential UK review of breast screening.**
 - 3 Gotzsche PC, Jorgensen KJ. Screening for breast cancer with mammography. *Cochrane Database Syst. Rev.* 26, CD001877 (2013).
 - 4 Puliti D, Duffy SW, Miccinesi G *et al.* Euroscreen Working Group. Overdiagnosis in mammographic screening for breast cancer in Europe: a literature review. *J. Med. Screen* 19(Suppl. 1), S42–S56 (2012).
 - 5 Bleyer A, Welch HG. Effect of three decades of screening mammography on breast-cancer incidence. *N. Engl. J. Med.* 367, 1998–2005 (2012).
 - 6 Kalager M, Zelen M, Langmark F, Adami O. Effect of screening mammography on breast-cancer mortality in Norway. *N. Engl. J. Med.* 363, 1203–1210 (2010).
 - 7 Duffy SW, Smith RA. More on screening mammography. *N. Engl. J. Med.* 364, 283 (2011).
 - 8 Kopans DB. Point: The New England Journal of Medicine Article suggesting overdiagnosis from mammography screening is scientifically incorrect and should be withdrawn. *J. Am. Coll. Radiol.* 10(5), 317–319 (2013).
 - 9 Nelson HD, Tyne K, Naik A, Bougatsos C, Chan BK, Humphrey L. Screening for breast cancer: an update for the US Preventive Services Task Force. *Ann. Intern. Med.* 151, 727–737 (2009).
 - 10 Hellquist BN, Duffy SW, Abdsaleh S *et al.* Effectiveness of population-based breast screening with mammography for women ages 40–49 years. *Cancer* 117, 714–722 (2011).
 - 11 Berry DA, Cronin KA, Plevritis SK *et al.* Effect of screening and adjuvant therapy on mortality from breast cancer. *N. Engl. J. Med.* 353, 1784–1792 (2005).
 - 12 Pashayan N, Duffy SW, Chowdhury S *et al.* Polygenic susceptibility to prostate and breast cancer: implications for personalised screening. *Br. J. Cancer* 104, 1656–1663 (2011).
 - 13 Duffy SW, Yen AM, Chen TH *et al.* Long-term benefits of breast screening. *Breast Cancer Manag.* 1, 31–38 (2012).
 - 14 EUROSREEN working group. Summary of the evidence of breast cancer service screening outcomes in Europe and first estimate of the benefit and harm balance sheet. *J. Med. Screen* 19(Suppl. 1), S5–S13 (2012).
 - **Review of outcomes of breast cancer service mammography screening in Europe.**
 - 15 Sasieni PD. Outcomes of screening to prevent cancer: think of screening as insurance. *BMJ* 327, 50 (2003).
 - 16 Beral V, Alexander M, Duffy S *et al.* The number of women who would need to be screened regularly by mammography to prevent one death from breast cancer. *J. Med. Screen* 18, 210–212 (2011).
 - 17 Autier P, Boniol M, La Vecchia C *et al.* Disparities in breast cancer mortality trends between 30 European countries: Retrospective analysis of WHO mortality database. *BMJ* 341, c3620 (2010).
 - 18 Broeders M, Moss S, Nyström L *et al.* The impact of mammographic screening on breast cancer mortality in Europe: a review of observational studies. *J. Med. Screen* 19(Suppl. 1), S14–S25 (2012).

- 19 Gabe R, Duffy SW. Evaluation of service screening mammography in practice: the impact on breast cancer mortality. *Ann. Oncol.* 16(Suppl. 2), ii153–ii162 (2005).
 - 20 Sasieni P, Castanon A, Cuzick J. Effectiveness of cervical screening with age: population based case-control study of prospectively recorded data. *Br. Med. J.* 339, b2968 (2009).
 - 21 Schultz H, Bohlius J, Skoetz N *et al.* Chemotherapy plus Rituximab versus chemotherapy alone for B-cell non-Hodgkin's lymphoma. *Cochrane Database Syst. Rev.* (4), CD003805 (2007).
 - 22 Jørgensen KJ, Gøtzsche PC. Overdiagnosis in publicly organised mammography screening programmes: systematic review of incidence trends. *BMJ* 339, b2587 (2009).
 - 23 Kalager M, Adami HO, Bretthauer M, Tamimi RM. Overdiagnosis of invasive breast cancer due to mammography screening: results from the Norwegian screening program. *Ann. Intern. Med.* 156, 491–499 (2012).
 - 24 Duffy SW, Parmar D. Overdiagnosis in breast cancer screening: the importance of long term observation and lead time. *Breast Cancer Res.* 15, R41 (2013).
 - 25 Duffy SW, Tabar L, Olsen AH *et al.* Cancer mortality in the 50–69 year age group before and after screening. *J. Med. Screen* 17, 159–160 (2010).
 - 26 Organising Committee and Collaborators, Falun meeting. Breast cancer screening with mammography in women aged 40–49 years. *Int. J. Cancer* 68, 693–699 (1996).
 - 27 Kerlikowske K. Screening mammography in women less than age 50 years. *Curr. Opin. Obstet. Gynecol.* 24, 38–43 (2012).
 - 28 Smith RA, Duffy SW, Gabe R, Tabar L, Yen AM, Chen TH. The randomized trials of breast cancer screening: what have we learned? *Radiol. Clin. North Am.* 42, 793–806 (2004).
 - 29 Menon U, Gentry-Maharaj A, Hallett R *et al.* Sensitivity and specificity of multimodal and ultrasound screening for ovarian cancer, and stage distribution of detected cancers: results of the prevalence screen of the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS). *Lancet Oncol.* 10, 327–340 (2009).
 - 30 Office for National Statistics. *Cancer Statistics: Registrations 2005*. HMSO, London, UK (2008).
 - 31 Smith RA, Kerlikowske K, Miglioretti DL, Kalager M. Clinical decisions. Mammography screening for breast cancer. *N. Engl. J. Med.* 367, e31 (2012).
 - 32 Assi V, Warwick J, Cuzick J, Duffy SW. Clinical and epidemiological issues in mammographic density. *Nat. Rev. Clin. Oncol.* 9, 33–40 (2011).
 - 33 Berg WA, Zhang Z, Lehrer D *et al.* Detection of breast cancer with addition of annual screening ultrasound or a single screening MRI to mammography in women with elevated breast cancer risk. *JAMA* 307, 1394–1404 (2012).
- **Key paper on varying imaging technology for specific groups.**
- 34 Lander MR, Tabar L. Autmoated 3-D breast ultrasound as a promising adjunctive screening tool for examining dense breast tissue. *Semin. Roentgenol.* 46, 302–308 (2011).
- **Introduction to new technology suitable for dense breasts.**
- 35 Maurice A, Evans DG, Affen J, Greenhalgh R, Duffy SW, Howell A. Surveillance of women at increased risk of breast cancer using mammography and clinical breast examination: further evidence of benefit. *Int. J. Cancer* 131, 417–425 (2012).
 - 36 Darabi H, Czene K, Zhao W, Liu J, Hall P, Humphreys K. Breast cancer risk prediction and individualised screening based on common genetic variation and breast density measurement. *Breast Cancer Res.* 14, R25 (2012).
 - 37 Chiu SY, Duffy SW, Yen AM, Tabar L, Smith RA, Chen HH. Effect of baseline breast density on breast cancer incidence, stage, mortality, and screening parameters: 25-year follow-up of a Swedish mammographic screening. *Cancer Epidemiol. Biomarkers Prev.* 19, 1219–1228 (2010).
 - 38 Schousboe JT, Kerlikowske K, Loh A, Cummings SR. Personalizing mammography by breast density and other risk factors for breast cancer: analysis of health benefits and cost-effectiveness. *Ann. Intern. Med.* 155, 10–20 (2011).
 - 39 Mandelblatt JS, Stout N, Trentham-Dietz A. To screen or not to screen women in their 40s for breast cancer: is personalized risk-based screening the answer? *Ann. Intern. Med.* 155, 58–60 (2011).
 - 40 Saslow D, Boates C, Burke W *et al.* American Cancer Society guidelines for breast screening with MRI as an adjunct to mammography. *CA Cancer J. Clin.* 57, 90–104 (2007).
 - 41 Malmgren JA, Parikh J, Atwood MK, Kaplan HG. Impact of mammography detection on the course of breast cancer in women aged 40–49 years. *Radiology* 262, 797–806 (2012).
 - 42 Duffy SW, Tabar L, Olsen AH *et al.* Absolute numbers of lives saved and overdiagnosis in breast cancer screening, from a randomized trial and from the breast screening programme in England. *J. Med. Screen* 17, 25–30 (2010).
 - 43 Jensen AR, Ewertz M, Cold S, Storm HH, Overgaard J. Time trends and regional differences in registration, stage distribution, surgical management and survival of breast cancer in Denmark. *Eur. J. Cancer* 39, 1783–1793 (2003).
 - 44 Swedish Organised Service Screening Evaluation Group. Effect of mammographic service screening on stage at presentation of breast cancers in Sweden. *Cancer* 109, 2205–2212 (2007).
- **Website**
- 101 Cancer research UK. Non-Hodgkin lymphoma mortality statistics. www.cancerresearchuk.org/cancer-info/cancerstats/types/nhl/mortality/uk-nonhodgkin-lymphoma-mortality-statistics